IOP Publishing | Bureau International des Poids et Mesures Metrologia

Metrologia 58 (2021) 055011 (7pp) https://doi.org/10.1088/1681-7575/ac1e03

A bilateral comparison of NIST and PTB laser power standards for scale realization confidence by gravitational wave observatories

Matthew Spidell[1,*], John Lehman[1], Marco López[2], Holger Lecher[2], Stefan Kück[2], Dripta Bhattacharjee[3], Yannick Lecoeuche[3] and Richard Savage[3]

[1] Sources and Detectors Group, US National Institute of Standards and Technology, Boulder, Colorado, United States of America
[2] Optics Division, Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
[3] LIGO Hanford Observatory, Richland, Washington, United States of America

E-mail: Matthew.Spidell@NIST.GOV

Received 9 June 2021, revised 29 July 2021
Accepted for publication 16 August 2021
Published 13 September 2021

Abstract

The gravitational wave (GW) observatories calibrate interferometer displacement using photon momentum, with laser power serving as the measurand. These observatories are traceable to the International System of Units through a primary standard maintained by the US's National Metrology Institute (NMI), the National Institute of Standards and Technology (NIST). The bilateral degree of equivalence of laser power measurements for various NMIs indicated in the 2010 EUROMET.PR-S2 supplementary comparison reveals scale realization uncertainty unacceptably large for GW event parameterization. We offer here an analysis to identify the source of the discrepancy between the Physikalisch-Technische Bundesanstalt (PTB) and NIST results. Using an improved transfer standard in a bilateral comparison, with representatives of the Laser Interferometer Gravitational-Wave Observatory (LIGO) receiving results prior to their comparison, NIST and PTB demonstrated a degree of equivalence of -0.15% with an uncertainty of 0.95% ($k = 2$) for combined 100 mW and 300 mW comparison results.

Keywords: gravitational wave observatory, photon calibrator, optical power calibration

(Some figures may appear in colour only in the online journal)

1. Introduction

Gravitational wave (GW) observatories currently calibrate the displacement of mirrored test masses at the ends of the interferometer arms by means of photon momentum [1–6]. In these interferometers, the absolute displacement of the mirror test masses is proportional to the absolute power ($\sim$1 W) of the optical force actuator. Consequently, the uncertainty in the interferometer displacement is directly proportional to the uncertainty in the power reflecting from the mirrored test mass, which also impacts the extraction of GW source parameters such as distance and sky location [7, 8]. Therefore, to achieve a maximum degree of reliability in the measurement of optical power, the calibration traceability of the optical power, and thus of the force, is organized through the photon calibrator (Pcal) program at LIGO [4, 9].

Between 2005 and 2007, several National Metrology Institutes (NMIs) undertook a supplementary comparison (2010

* Author to whom any correspondence should be addressed.

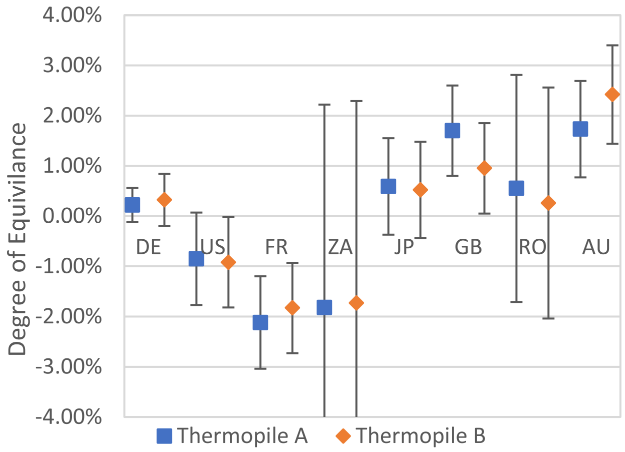

Figure 1. DoE and expanded uncertainty ($k = 2$) for the measurement of 1 W at 1064 nm. From the EUROMET.PR-S2 report [10] (ZA uncertainty bars truncated). Apparent trends in the results do not represent a chronological drift of the transfer standards as they were periodically re-evaluated at PTB as described in [10].

EUROMET.PR-S2) of their respective optical power calibration capability reported by Kück *et al* in [10]. This was accomplished by measuring the responsivity of two thermopile transfer standard detectors operating at a power range of 1 W. The results of this comparison are shown in figure 1. As can be seen, the differences in the determination of the responsivity of the thermopile detector are within a few percent for the participating NMIs. While this comparison generally demonstrated agreement among various NMIs, in some cases the degree of equivalence (DoE) between the NMI's reported value and the consensus value (i.e. unilateral DoE) approached or exceeded the uncertainty level of some of the participants. This concerned the GW research community and generated commentary suggesting that the 'standard for optical power varied by a few percent among different countries' metrology institutes' [6].

For some NMIs, their difference from the reference value (a weighted mean of participating NMI measurements) falls within their uncertainty. The responsivity determined by PTB Germany (DE), NMISA South Africa (ZA), NMIJ Japan (JP) and INM Romania (RO) are therefore consistent with the reference value. However, at least some of the NMIs may have underestimated their measurement uncertainty. The responsivity of the thermopile determined by NIST United States of America (US), LNE France (FR), NPL Great Britain (GB), and NMI Australia (AU) are either inconsistent or only partly consistent with the reference value.

After a discussion of the relevance of the EUROMET.PR-S2 supplementary comparison during the GW Metrology Workshop in 2019 [11], a bilateral comparison with a Pcal laser power transfer standard (described in the results section) was proposed between PTB and NIST due to the low uncertainty that both laboratories provide. The Pcal laser power transfer standard was chosen because LIGO's Pcal program previously demonstrated that this instrument presents low risk of measurement inequivalence when operated with differing injection formats [4, 9]. In this work we describe the outcome of the bilateral comparison between PTB and NIST for the optical powers of 100 mW and 300 mW at the wavelength of 1047 nm. The power was limited to 300 mW by available laser source and the 1047 nm wavelength matches that of the PCAL system. We compare this with the EUROMET.PR-S2 supplementary comparison at an optical power of 1 W and a wavelength of 1064 nm.

2. Realization of the optical radiant flux scale

Cryogenic absolute radiometers permit a detector-based realization of the scale for optical radiant flux ('power' or 'laser power') [12] with a typical uncertainty below 0.05% for an incident power of approximately 100 μW [13, 14]. To achieve detector-based measurements at higher powers, an NMI typically employs a cryogenic radiometer to calibrate a transfer standard which is then used to calibrate other transfer standards at progressively higher powers. With each stage transferring the calibration at higher power, an additional uncertainty is introduced. The uncertainty achieved by PTB through this chain during the 2010 and the present comparison is 0.2%. An alternative system using an electrically-calibrated primary standard calorimeter is employed at NIST for direct representation of the optical power at the 200 mW level [15, 16] (with a calibrated beam splitter allowing calibrations at higher power). The uncertainty achieved through this technique was 0.86% during the 2010 comparison and 0.84% during the present comparison. While great care is taken to prevent the introduction of systematic error and identify uncertainty contributions, the introduction of unknown systematic errors in a measurement process is still possible. Therefore, intercomparison of laboratories is required to identify discrepancies in power scale representation, which may indicate unidentified systematic errors.

3. Previous international comparison at the 1W level

The EUROMET.PR-S2 supplementary comparison results [10] show the bilateral DoE for calibrations of the thermopile instruments performed by PTB and NIST with an optical power of 1 W and 1064 nm wavelength slightly exceeded the $k = 2$ bilateral uncertainty level as listed in table 1. The low DoE assessed for PTB and NIST at 514 nm contrasted with larger DoE assessed for 1064 nm measurements suggests an unaccounted systematic error during 1064 nm measurements.

With both laboratories making use of cryogenic radiometry to validate the spectral responsivity, an alignment deficiency presents a favorable explanation for the apparently correlated (systematic) inequivalence contribution. While measurement stochastics cannot be ruled out owing to the magnitude of the bilateral uncertainty, the likelihood of repeated discrepancy with magnitude approximately $k = 2$, with uncorrelated inequivalence contributions, is much less than 1%. As the transfer standards were evaluated at PTB multiple times during the international comparison activity and demonstrated repeatability much less than the DoE, a misalignment

Table 1. EUROMET.PR-S2 supplementary comparison bilateral DoE for PTB and NIST for thermopile-based instruments A and B.

Thermopile Unit	Power W	DoE %	Uncertainty ($k = 2$) %
	1064 nm wavelength		
A	1	1.07	1.00
B	1	1.24	1.07
	514 nm wavelength		
A	1	−0.20%	0.98
B	1	0.01	0.95

most-likely occurred during the 1064 nm calibrations at the NIST.

4. Thermopile transfer standard inequivalence contributions

Thermopile-type power detectors typically consist of an optical absorber coupled to a multi-junction thermopile, which is coupled to a thermal reservoir. The temperature of the thermal reservoir affects the responsivity of the thermopile. During the EUROMET.PR-S2 supplementary comparison, NIST used an injection time of 200 s, while PTB used an injection time of 600 s. The longer injection period at PTB heats the thermal reservoir more than the shorter injection at NIST. A thermopile similar to those used in the EUROMET.PR-S2 supplementary comparison was used to model the difference in responsivity attributable to the differing injection times used at PTB and NIST. Our model predicts an inequivalence of approximately 0.12% when responsivity is averaged over the evaluation period used by PTB versus the shorter integration period used by NIST. To be clear, this is not an inequivalence in the scale representation at PTB versus NIST, it is the inequivalence in thermopile responsivity attributable to differing injection times. The decrease in relative responsivity of a typical thermopile is plotted in figure 2 with injection timing differences between NMIs annotated.

Spatial non-uniformity may be evaluated by scanning a laser across the surface of the detector, compensating for source drift by repeatedly referencing the peripheral responsivity against the responsivity at the center of the instrument. The details of the method are described in [17]. Spatial non-uniformity in instrument responsivity is attributable to a variety of mechanical or optical non-uniformities in device construction. The resulting inequivalence is device- and operator-specific which precludes an equitable treatment of the uncertainty contribution assessment for the EUROMET.PR-S2 supplementary comparison at this time. However, spatial non-uniformity for a thermopile similar to those used during the EUROMET.PR-S2 supplementary comparison is depicted in figure 3. It follows that an alignment discrepancy of 2 mm can readily yield a responsivity variation exceeding 1%.

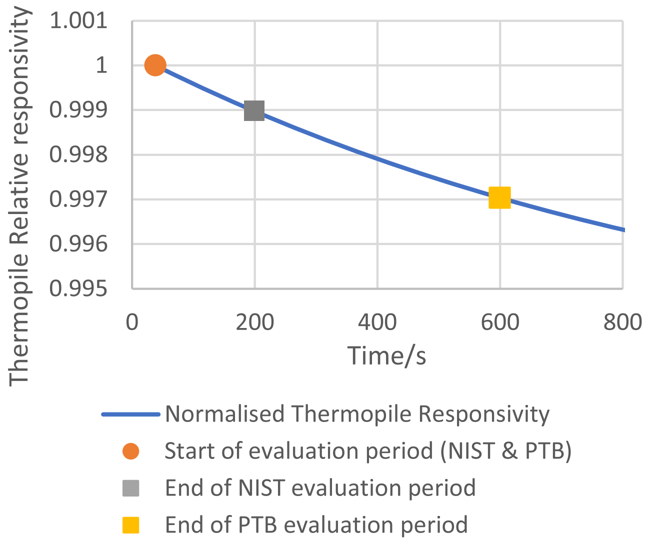

Figure 2. Typical thermopile relative responsivity versus injection time with 1 W of incident power. Time is measured from the start of the injection.

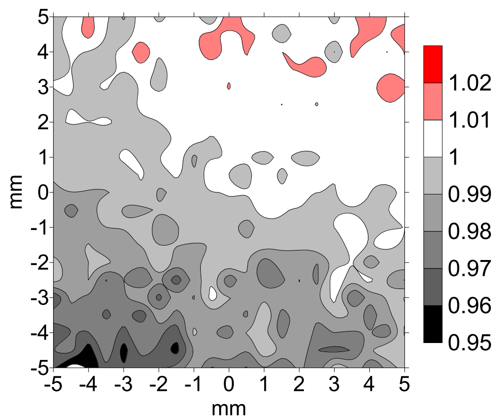

Figure 3. Example thermopile spatial non-uniformity. Evaluated with a resolution of 0.5 mm over a 10 mm square. Measurements were centered within the detector surface.

5. An integrating sphere transfer standard for a NIST and PTB bilateral comparison

In order to improve on the thermal and spatial non-uniformities of the instrument used for the 2010 EUROMET.PR-S2 supplementary comparison, our current comparison used an integrating-sphere-based transfer standard. Integrating spheres offer significant reduction in inequivalence, attributable to laser induced heating of the instrument, and excellent spatial uniformity. The instrument used in this comparison is a 100 mm diameter integrating sphere with an aluminum outer shell, a sintered PTFE inner shell, a 25 mm diameter entrance aperture, and a 12.7 mm diameter detector port. It is similar to the 'Gold Standard' described in [7] but has not been calibrated at an NMI previously. A baffle limits the field of view to the portion of the wall not directly illuminated by the laser. The FOV is further limited by the detector's recessed attachment point.

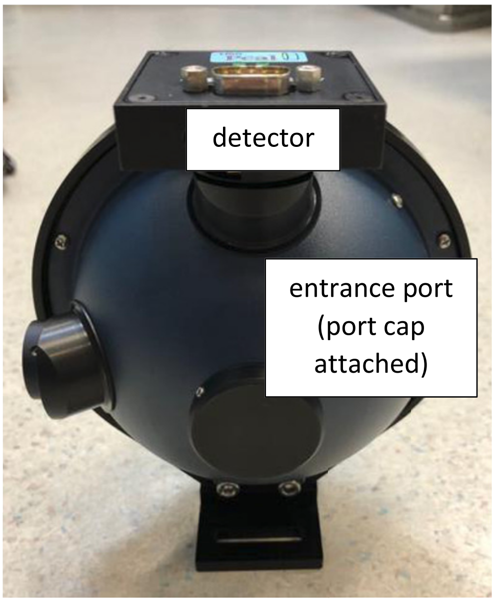

Figure 4. Image of the transfer standard assembly (front view with port cap attached).

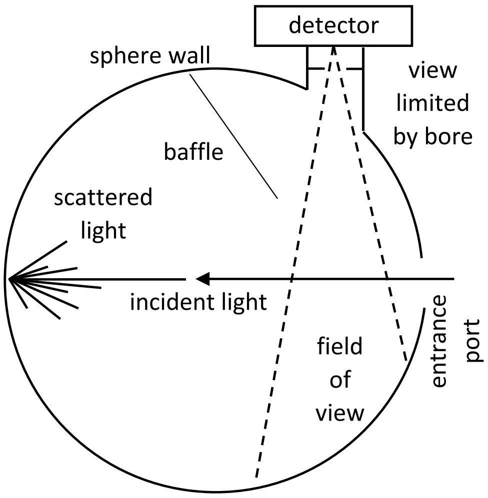

Figure 5. Diagram of the transfer standard assembly (side view).

The detector consists of an indium gallium arsenide (InGaAs) photodiode coupled as shown in figures 4 and 5.

Such instruments have been successfully employed by LIGO's Pcal program since 2006 as transfer standards demonstrating long-term stability under 0.1% and the instrument's thermal coefficient of responsivity in the present configuration is 0.096% K^{-1} [9, 18]. Principal contributions to thermal coefficient of responsivity are the coefficient of thermal expansion of the sintered PTFE interior shell and the thermal coefficient of responsivity of the photodiode. Due to the greater heat dissipation provided by the large area integrating sphere, lower power used in the comparison, and the lower absorptivity of the instrument the thermal rise of the instrument during measurements was negligible. Moderate differences in laboratory ambient temperature were compensated in reporting results.

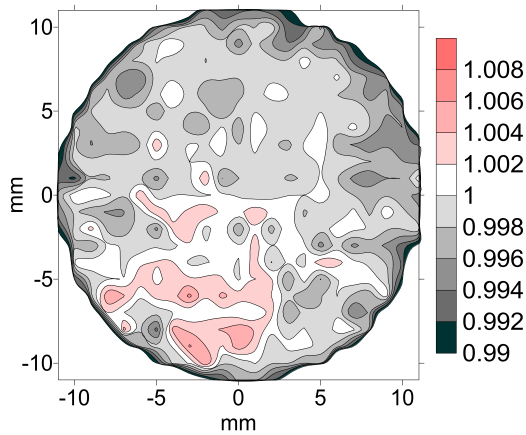

Figure 6. Spatial non-uniformity of the integrating sphere transfer standard assembly shown in figures 2 and 3. Evaluated with a resolution of 0.5 mm over a 20 mm square.

As depicted in figure 6 the spatial uniformity of a typical 100 mm integrating sphere, with baffles and aperture configured for laser power measurement, is improved an order-of-magnitude beyond that of a thermopile. Removing all objects from the area surrounding the integrating sphere entrance aperture is critical. During prior work we noted that scattered light from nearby objects, such as apertures, caused inequivalence exceeding 0.5%. Empirically, we determined, no object should be located closer than 15 cm to the 2.5 cm aperture, with greater distance preferred.

6. Methodology

In the current comparison, the responsivity of the transfer standard was determined first at NIST, then at PTB and finally again at NIST. The methodology used at NIST is described in [19] while the methodology used at PTB is described in [20]. These methodologies are the same used for typical laser power meter calibration provided by NIST or PTB and meet the documentation requirements of ISO 17025:2017. Results were reported to LIGO which disseminated them, so that the results of a given laboratory had no influence upon the results of the other. Both laboratories used the same light source; a 1047 nm laser which traveled with the transfer standard and performed calibrations at 100 mW and 300 mW. An uncalibrated temperature sensor (onboard transducer) with a nominal sensitivity of 0.01 V/°C was co-located with the transfer standard's photodiode and was used for precise relative temperature correction. Laboratory temperatures and corresponding onboard transducer voltages are shown in table 2.

7. Results

Measurement results, adjusted to a common onboard temperature transducer voltage of 3.0089 V (corresponding with PTB laboratory conditions), are shown in table 3 and plotted in figure 7. (Results were adjusted by using the 0.096% K^{-1} coefficient of responsivity and reported laboratory temperatures). Close agreement in measurements at NIST prior to and after

Table 2. Average temperature of the laboratory and the average voltage of the transfer standard's temperature transducer.

NMI Units	Laboratory temperature °C	Limits ±°C	Onboard transducer V
NIST (Feb 2020)	20.6	1	2.9974
PTB (May 2020)	21.5	0.5	3.0089
NIST (Oct 2020)	21.0	1	3.0073

Table 3. Transfer standard responsivity (R) adjusted for a consensus onboard temperature sensor voltage 3.0089 V.

NMI Units	Power mW	R V W^{-1}	$u(R)$ $(k = 1)$ %
NIST (Feb 2020)	100 mW	8.189	0.42
	300 mW	8.202	0.42
PTB (May 2020)	100 mW	8.190	0.10
	300 mW	8.186	0.10
NIST (Oct 2020)	100 mW	8.202	0.42
	300 mW	8.207	0.42

measurement at PTB implies that the instrument endured shipping without damage. The instrument showed no significant difference in responsivity at 100 mW and 300 mW, suggesting that the non-linearity falls below the measurement uncertainty.

8. Analysis

A comparison of laboratory results follows the procedures described by the International Bureau of Weights and Measures (BIPM) in [21]. A comparison was performed for the 100 mw and 300 mw responsivities listed in table 3. We also compare a composite responsivity, which is the average of the 100 mw and 300 mw responsivity determined by each laboratory.

For each power level, the relative difference between the laboratory measurement and the reference responsivity is determined by

$$\Delta_l = \frac{R_l}{RR} - 1 \quad (1)$$

with $R_l = R_{\mathrm{PTB}}$ or $R_l = \left(R_{\mathrm{NIST,Feb\ 2020}} + R_{\mathrm{NIST,Oct\ 2020}}\right)/2$, respectively. Because of their lower uncertainty, the responsivity values provided by PTB have been chosen to serve as the reference responsivity RR (i.e. $RR = R_{\mathrm{PTB}}$).

For each power level, the relative uncertainty is determined by

$$u\left(\Delta_l\right) = \sqrt{u^2\left(R_l\right) + {u_{\mathrm{ts}}}^2}, \quad (2)$$

where R_l is the responsivity at each NMI, and $u\left(R_l\right)$ is the NMI's standard measurement uncertainty for the comparison. The uncertainty in the transfer standard u_{ts} is 0.1% as demonstrated by LIGO's Pcal group [4].

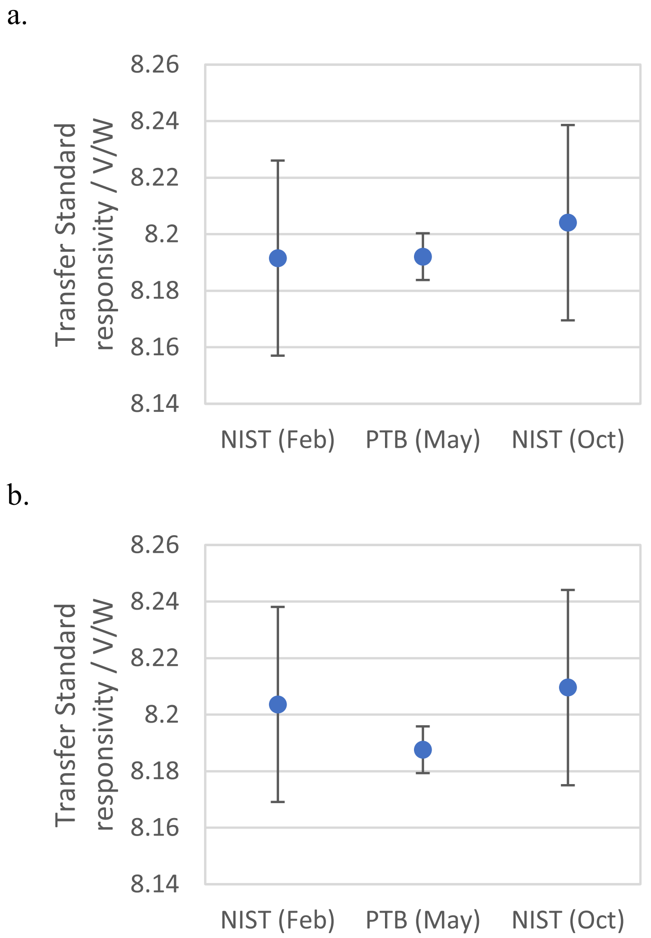

Figure 7. Absolute responsivity of the transfer standard as evaluated at NIST and PTB using powers of 100 mW (a) and 300 mW (b) at 1047 nm temperature-adjusted to consensus onboard temperature sensor voltage of 3.0089 V. Error bars show the combined standard uncertainty level 'u' reported by each laboratory.

The uncertainty weighting for each laboratory's results was determined by

$$w_l = u^{-2}\left(\Delta_l\right) \Big/ \sum_{l=1}^{2} u^{-2}\left(\Delta_l\right) \quad (3)$$

which is summarized for each laboratory and power level in table 4.

Consensus responsivity is determined by

$$CR = \left(1 + \Delta_{RR}\right) RR. \quad (4)$$

Where Δ_{RR} is determined by

$$\Delta_{RR} = \sum_{l=1}^{2} w_l \Delta_l \quad (5)$$

and the corresponding uncertainties are:

$$u\left(\Delta_{RR}\right) = 1 \Big/ \sqrt{\sum_{i=1}^{2} u^{-2}\left(\Delta_l\right)} \quad (6)$$

$$u\left(CR\right) = \sqrt{\left(\Delta_{RR} u\left(\Delta_{RR}\right)\right)^2 + u(RR)^2} \quad (7)$$

Table 4. Establishing participant laboratory weight w_l. u_{TS}: transfer standard uncertainty ($k = 1$), when transported between laboratories, Δ_l: relative responsivity difference between participant and pilot laboratory, $u(\Delta_l)$: standard uncertainty of Δ_l, w_l: weight ascribed to the responsivity reported by the participant laboratory.

Participant Units	Power mW	u_{TS} %	Δ_l %	$u(\Delta_l)$ %	w_l
PTB			0.00	0.14	0.903
NIST	100		0.07	0.43	0.097
PTB			0.00	0.14	0.903
NIST	300		0.23	0.43	0.097
PTB			0.00	0.14	0.903
NIST	Composite	0.1	0.15	0.43	0.097

Table 5. Establishing consensus responsivity *CR* for an onboard temperature sensor voltage of 3.0089 V. *RR*: reference responsivity, Δ_{RR}: relative difference between *RR* and *CR*, $u(\Delta_{RR})$: standard uncertainty of Δ_{RR}.

Power Units	*RR* V W^{-1}	Δ_{RR} %	$u(\Delta_{RR})$ %	*CR* V W^{-1}	$u(CR)$ %
100 mW	8.190	0.007	0.14	8.192	0.101
300 mW	8.186	0.023	0.14	8.189	0.101
Composite	8.188	0.015	0.14	8.189	0.101

Table 6. Chi-squared values and consistency check.

Power	χ^2_{obs}	$\chi^2_{0.05}$	Consistency
100 mW	0.024	3.841	Satisfied
300 mW	0.26	3.841	Satisfied
Composite	0.11	3.841	Satisfied

with results presented in table 5.

Using the guidelines in [21] describing procedures for a consistency check using chi-squared values, the uncertainty of the consensus responsivity is validated when $\nu = 2$ $\chi^2_{0.05}$ exceeds the observed $\chi^2{}_{\mathrm{obs}}$ value which is determined by

$$\chi^2_{\mathrm{obs}} = \sum_{l=1}^{2} (\Delta_l - \Delta_{RR})^2 / u^2 (\Delta_l) \tag{8}$$

with results presented in table 6.

The DoE is determined by

$$\mathrm{DoE} = \Delta_{\mathrm{PTB}} - \Delta_{\mathrm{NIST}}, \tag{9}$$

where PTB and NIST responsivities are substituted into the expression for Δ_l defined in equation (1). Uncertainty is defined by

$$U(\mathrm{DoE}) = \sqrt{u^2 (\Delta_{\mathrm{PTB}}) + u^2 (\Delta_{\mathrm{NIST}})} \tag{10}$$

with results presented in table 7.

The relative difference in measured responsivity R_l from consensus responsivity *CR* for each power is determined by

$$\Delta_{R,l} = R_l / CR \tag{11}$$

Table 7. Present bilateral DoE for PTB and NIST.

Power Units	DoE %	U ($k = 2$) %
100 mW	−0.07	0.91
300 mW	−0.23	0.91
Composite	−0.15	0.87

Table 8. $\Delta_{R,l}$: variation in laboratory responsivity from consensus responsivity, χ: ratio of $\Delta_{R,l}$ to $u(R_l)$, and p_{Δ_R}: probability of observing variation of this magnitude. Lower probabilities are preferred.

Participant Units	Power	$\Delta_{R,l}$ %	χ %	$p_{\Delta_{R,l}}$ %
PTB		−0.007	7	5
NIST	100 mW	0.063	15	12
PTB		−0.023	22	18
NIST	300 mW	0.209	50	38
PTB		−0.015	15	12
NIST	Composite	0.136	32	25

and the ratio of $\Delta_{R,l}$ to the laboratory's measurement uncertainty $u(R_l)$ for each power is determined by

$$\chi = \Delta_{R,l} / u (R_l) . \tag{12}$$

Using the expression for χ and the cumulative probability distribution function for the normal distribution we determine the probability of observing closer agreement. Results are listed in table 8.

9. Discussion

The results of this study show bilateral DoE much lower than the comparison uncertainty, see table 7. This suggests that the uncorrelated systematic error falls well below the uncertainty level reported. Due to the low uncertainty achieved by PTB, which establishes the reference responsivity, the difference between the consensus responsivity and reference responsivity is trivial for both power levels and the composite result. The uncertainty in consensus responsivity prevents a firm conclusion regarding linearity of the transfer standard, however, the close agreement in responsivity at 100 mW and 300 mW suggests that the non-linearity is negligible over this power range. Because the non-linearity is much smaller than the uncertainty, the composite responsivity provides the best determination of instrument responsivity. The low probability of closer agreement for most results listed in table 8 provides another perspective to validate the consistency of results.

10. Conclusion

The composite DoE of −0.15% with an uncertainty of 0.87% (for $k = 2$) obtained in this bilateral comparison validates the scale representation realized by both laboratories and the use of photodiode-coupled integrating spheres as transfer standards. LIGO's representation of the optical power is accurate to the stated calibration uncertainty and confidence interval. The optical power scale realization by NIST and PTB is adequate, enabling Pcal-induced displacement uncertainty of 0.82% ($k = 2$) for the recently-completed third observing run of the observatory [9]. As GW observatories reduce other uncertainty contributions to mirror displacement measurements, further reduction in optical power scale uncertainty may be required. NIST is developing a new standard with lower expected uncertainty described in [22] which is expected to offer transfer standard calibration uncertainty below 0.15%. This will support future inter-comparisons utilizing photodiode-coupled integrating spheres as transfer standards. It also suggests an opportunity for co-locating room temperature primary radiometric standards at GW observatories to achieve the lowest achievable power measurement uncertainty.

Acknowledgments

This is a work of the US Federal Government and therefore exempt from US copyright.

ORCID iDs

Matthew Spidell https://orcid.org/0000-0002-5513-8520
John Lehman https://orcid.org/0000-0003-4729-7123
Marco López https://orcid.org/0000-0002-0706-1418
Stefan Kück https://orcid.org/0000-0003-2188-483X
Dripta Bhattacharjee https://orcid.org/0000-0001-6623-9506
Richard Savage https://orcid.org/0000-0003-3317-1036

References

[1] Abernathy M R *et al* 2015 Advanced LIGO *Class. Quantum Grav.* **32** 074001
[2] Acernese F *et al* 2015 Advanced Virgo: a second-generation interferometric gravitational wave detector *Class. Quantum Grav.* **32** 024001
[3] Aso Y *et al* 2013 Interferometer design of the KAGRA gravitational wave detector *Phys. Rev.* D **88** 043007
[4] Karki S *et al* 2016 The Advanced LIGO photon calibrators *Rev. Sci. Instrum.* **87** 114503
[5] Estevez D *et al* 2021 The Advanced Virgo photon calibrators *Class. Quantum Grav.* **38** 075007
[6] Inoue Y *et al* 2018 Improving the absolute accuracy of the gravitational wave detectors by combining the photon pressure and gravity field calibrators *Phys. Rev.* D **98** 022005
[7] Abbott B *et al* 2017 Calibration of the Advanced LIGO detectors for the discovery of the binary black-hole merger GW150914 *Phys. Rev.* D **95** 062003
[8] Abbott R *et al* 2021 GWTC-2: compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run *Phys. Rev.* X **11** 021053
[9] Bhattacharjee D *et al* 2021 Fiducial displacements with improved accuracy for the global network of gravitational wave detectors *Class. Quantum Grav.* **38** 15009
[10] Kück S 2010 Final report on EUROMET comparison EUROMET.PR-S2 (Project No. 156): responsivity of detectors for radiant power of lasers *Metrologia* **47** 02003
[11] Lehman J 2019 *GW (Gravitational Wave) Metrology Workshop Executive Summary* https://dcc.ligo.org/LIGO-L1900166/public
[12] Martin J, Fox N and Key P 1985 A cryogenic radiometer for absolute radiometric measurements *Metrologia* **21** 147–55
[13] Livigni D 2003 *High Accuracy Laser Power Energy Meter Calibration Service* (Boulder, CO: NIST) https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=30884
[14] Werner L, Fischer J, Johannsen U and Hartmann J 2000 Accurate determination of the spectral responsivity of silicon trap detectors between 238 nm and 1015 nm using a laser-based cryogenic radiometer *Metrologia* **37** 279–84
[15] West E, Case W, Rasmussen A L and Schmidt L 1972 A reference calorimeter for laser energy measurement *J. Res. Natl. Bur. Stand.* **1** 13–26
[16] Spidell M and Vaskuri A 2021 Optical power scale realization by laser calorimeter after 45 years of operation *NIST J. Res.* **126** 126011
[17] Livigni D and Li X 1996 Spatial uniformity of optical detector responsivity *NCSL Conf. Proc.*
[18] Lecoeuche Y, Karki S and Savage R 2020 Pcal WSS temperature coefficient https://dcc.ligo.org/LIGO-T2100263/public
[19] Hadler J A, Cromer C L and Lehman J H 2007 NIST measurement services: cw laser power and energy meter calibration service at NIST https://doi.org/10.6028/NIST.SP.250-75
[20] Stock K, Kück S and Brandt F 2003 Laserradiometrie *PTB-Mitteilungen* vol 113 p 361 www.ptb.de/cms/fileadmin/internet/publikationen/ptb_mitteilungen/mitt_pdf_vor_2007/2003/PTB-Mitteilungen_2003_4.pdf
[21] CCPR Working Group on Key Comparisons 2013 *Guidelines for CCPR Key Comparison Report Preparation* BIPM https://bipm.org/utils/common/pdf/CC/CCPR/CCPR-G2.pdf
[22] Vaskuri A, Stephens M, Tomlin N, Spidell M, Yung C, Walowitz A, Straatsma C, Harber D and Lehman J 2021 High-accuracy room temperature planar absolute radiometer based on vertically aligned carbon nanotubes *Opt. Express* **29** 22533–52